# 9 Lecture 9

## 9.1 Cramer's rule

**Theorem 9.1** (Cramer's rule)**.** *Let $A \in M_{n\times n}(F)$ be invertible and $b \in F^n$. Then the solution $x = (x_1, x_2, \ldots, x_n)^t$ to $Ax = b$ is given by:*

$$x_i = \frac{\det(A_i(b))}{\det(A)}, \quad \text{for } i = 1, 2, \ldots, n.$$

*Proof.* Let $x = (x_1, x_2, \ldots, x_n)^t$ be the unique solution to $Ax = b$. Then, $Ax = x_1v_1 + x_2v_2 + \cdots + x_nv_n = b$. Therefore $\det(A_i(b)) = \sum_{j=1}^n x_j \det(A_i(v_j)) = x_i \det(A)$ and hence $x_i = \frac{\det(A_i(b))}{\det(A)}$. ◻

**Example 9.2.** Consider the equation

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

If $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is invertible, then the unique solution to this equaiton is

$$x_1 = \frac{\det\begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}} = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}}, \quad x_2 = \frac{\det\begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}}{\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}} = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}}.$$

**Proposition 9.3.** *Let $M$ be an $n \times n$ matrix of the form*

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

*where $A \in M_{p\times p}(F)$, $B \in M_{p\times(n-p)}(F)$, $C \in M_{(n-p)\times(n-p)}(F)$ and $0$ is a matrix of all zeros. Show that*

$$\det(M) = \det(A)\det(C).$$

*Proof.* If $C$ is singular, then $\operatorname{rank} C < n - p$. Then the last $n - p$ rows of $\begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$ are linearly dependent. Hence $M$ is singular. So $\det(M) = 0 = \det(A)\det(C)$.

If $C$ is invertible, we consider the function $D(A) = \det(C)^{-1} \det \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$. Then $D(A)$ satisfies (i), (ii) in Definition 7.12 since det does. Moreover, by applying Proposition 8.14 repeatedly, we get

$$D(I) = \det(C)^{-1} \det \begin{pmatrix} I & B \\ 0 & C \end{pmatrix} = \det(C)^{-1} \det(C) = 1.$$

Thus by Theorem 8.1, we have $D(A) = \det(A)$. ◻

*Remark* 9.4. One can similarly show that

$$\det \begin{pmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_n \end{pmatrix} = \det(A_1)\ldots\det(A_n).$$

where $A_i$ are square matrices whose diagonal aligns with the diagonal of the whole matrix. This should be compared with Proposition 8.14.

## 9.2 Polynomials

This section is an introduction to polynomial theory. We only need the conclusions. For a proof, see [1, Chapter 9]. For a proof of the fundamental theorem of algebra see https://en.wikipedia.org/wiki/Fundamental_theorem_of_algebra.

**Definition 9.5.** Let $F$ be a field. The set of all polynomials in one variable $x$ with coefficients in $F$ is denoted by $F[x]$. Note that there is no restriction on the degree here. In other words, $F[x] = \cup_{n=1}^{\infty} \mathcal{P}_n(F)$. A general element of $F[x]$ has the form

$$f(x) = a_n x^n + \cdots + a_1 x + a_0,$$

where $a_0, \ldots, a_n \in F$, and $n \in \mathbb{N}$.

If $f(x) \in F[x]$ is nonzero, its **degree**, denoted $\deg(f)$, is the highest power of $x$ with nonzero coefficient. We define $\deg(0) = -\infty$ by convention.

A polynomial $f(x) \in F[x]$ is called **monic** if its leading coefficient is 1.

**Theorem 9.6.** *Given polynomials $f(x), g(x) \in F[x]$ with $g(x) \neq 0$, there exist unique polynomials $q(x), r(x) \in F[x]$ such that:*

$$f(x) = q(x)g(x) + r(x), \qquad \text{with } \deg(r) < \deg(g).$$

*This is called the **Euclidean division**.*

**Example 9.7.** $x^3 - 1 = (x^2 + x + 1)(x - 1)$.

**Definition 9.8.** Let $f(x), g(x) \in F[x]$. We say that $g(x)$ **divides** $f(x)$, written $g(x) \mid f(x)$, if there exists $q(x) \in F[x]$ such that $f(x) = q(x)g(x)$.

**Definition 9.9.** Let $f(x) \in F[x]$, and let $\alpha \in F$. We say that $\alpha$ is a **root** (or **zero**) of $f$ if $f(\alpha) = 0$.

**Proposition 9.10.** *$\alpha \in F$ is a root of $f(x)$ if and only if $(x - \alpha) \mid f(x)$ in $F[x]$.*

**Definition 9.11.** Let $p(x) \in F[x]$ be a nonzero polynomial of degree $n$. We say that $p(x)$ **splits in** $F$ if it can be written as a product of linear factors, that is,

$$p(x) = a(x - \lambda_1)(x - \lambda_2)\cdots(x - \lambda_n),$$

for some $a \neq 0$ and $\lambda_1, \dots, \lambda_n \in F$. One can also write it as

$$p(x) = a(x-\lambda_1)^{k_1}(x-\lambda_2)^{k_2} \cdots (x-\lambda_m)^{k_m}$$

where $\lambda_1, \dots, \lambda_m$ are *distinct* roots of $p$ and $k_i$ is the multiplicity of $\lambda_i$.

**Theorem 9.12** (Fundamental theorem of algebra). *Let $p(x) \in \mathbb{C}[x]$. Then $p(x)$ splits in $\mathbb{C}$.*

**Example 9.13.** Let $p(x) = ax^2 + bx + c$ where $a, b, c \in \mathbb{R}$. Then $p(x) = a(x-\lambda_1)(x-\lambda_2)$ in $\mathbb{C}[x]$ where

$$\lambda_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad \lambda_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

here you can take the square root of any real number since if $b^2 - 4ac \geq 0$ then it's the usual square root, if $b^2 - 4ac < 0$, then $\sqrt{b^2 - 4ac} = \sqrt{4ac - b^2}\sqrt{-1} = \sqrt{4ac - b^2}i$.

For example $x^2 + 1$ does not splits in $\mathbb{R}$ since it has no roots in $\mathbb{R}$. However, $x^2 + 1 = (x-i)(x+i)$ splits in $\mathbb{C}$ and the roots are $\pm i$.

Also $x^3 - 1 = (x-1)(x^2+x+1)$ does not splits in $\mathbb{R}$ since $x^2 + x + 1$ has no roots in $\mathbb{R}$. However, $x^3 - 1 = (x-1)(x-\lambda_1)(x-\lambda_2)$ in $\mathbb{C}$ where $\lambda_1 = \frac{-1+\sqrt{3}i}{2}$ and $\lambda_2 = \frac{-1-\sqrt{3}i}{2}$. Note that $\lambda_1^2 = (\frac{-1+\sqrt{3}i}{2})^2 = \frac{1-2\sqrt{3}i-3}{4} = \frac{-1-\sqrt{3}i}{2} = \lambda_2$. So the above formula is also written as $x^3 - 1 = (x-1)(x-\omega)(x-\omega^2)$ where $\omega = \lambda_1$.

**Definition 9.14.** Let $z = x + yi$ be a complex number $x, y \in \mathbb{R}$. The **complex conjugate** is the complex number $\bar{z} = x - yi$. $x$ is called the **real part** of $z$ and $y$ is called the **imaginary part** of $z$. We write $x = \operatorname{Re} z$ and $y = \operatorname{Im} z$. (Note that the imaginary part Im should not be confused with the image im.)

We have the following properties for complex conjugates: $\overline{zw} = \bar{z}\bar{w}$ and $\overline{z+w} = \bar{z}+\bar{w}$.

Let $A = (a_{ij}) \in M_{m\times n}(\mathbb{C})$. We define the **complex conjugate matrix** $\bar{A} = (\overline{a_{ij}}) \in M_{m\times n}(\mathbb{C})$ and $\operatorname{Re} A = (\operatorname{Re} a_{ij}) \in M_{m\times n}(\mathbb{R})$ and $\operatorname{Im} A = (\operatorname{Im} a_{ij}) \in M_{m\times n}(\mathbb{R})$. By the properties above, it's not hard to show that $\overline{AB} = \bar{A}\bar{B}$ and $\overline{A+C} = \bar{A} + \bar{C}$ for $A, C \in M_{m\times n}(\mathbb{C})$, $B \in M_{n\times p}(\mathbb{C})$.

**Example 9.15.** If $A = \begin{pmatrix} 1-i & 1+i \\ 1 & 1 \end{pmatrix}$, then $\bar{A} = \begin{pmatrix} 1+i & 1-i \\ 1 & 1 \end{pmatrix}$, $\operatorname{Re} A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ and $\operatorname{Im} A = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}$.

**Lemma 9.16.** *Suppose $p(x) \in \mathbb{R}[x]$ and we regard it as a polynomial in $\mathbb{C}[x]$. If $z \in \mathbb{C}$ is a root of $p$ i.e. $p(z) = 0$, then $\bar{z}$ is also a root of $p$ i.e. $p(\bar{z}) = 0$.*

*Proof.* Since $p(x) = a_n x^n + \cdots + a_0 \in \mathbb{R}[x]$, $\overline{a_i} = a_i$

$$p(\bar{z}) = a_n(\bar{z})^n + \cdots + a_1\bar{z} + a_0 = \overline{a_n z^n + \cdots + a_1 z + a_0} = \bar{0} = 0.$$

□

**Proposition 9.17.** *For any $p(x) \in F[x]$, there is a field extension $E$ of $F$ i.e. a field $E$ containing $F$, such that $p(x)$ splits in $E$. The smallest such field $E$ is called the* ***splitting field*** *of $p(x)$.*

## 9.3 Eigenvalue and eigenvectors

Let $T : V \to V$ be a linear transformation. We would like to find a basis such that $[T]_\beta^\beta$ is the simplest. The simplest possible matrix is the diagonal matrix $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$. Although this is not always achievable, if we do have such a basis $\beta = (v_1, \ldots, v_n)$, then the $v_i$ satisfies the special property that $T(v_i) = \lambda_i v_i$ for $1 \le i \le n$.

**Definition 9.18.** We say that $\lambda \in F$ is an **eigenvalue** of $T$ and $v \in V$ is an **eigenvector** corresponding to the eigenvalue $\lambda$ if $v \neq 0$ and $Tv = \lambda v$.

A linear transformation $T : V \to V$ is **diagonalizable** over $F$ if there is a basis $\beta$ of $V$ such that $[T]_\beta^\beta = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

The set of all eigenvalues of $T$ is called the **spectrum** of $T$ denoted by $\sigma(T)$.

In matrix language, $\lambda \in F$ is an **eigenvalue** of $A$ and $v \in F^n$ is an **eigenvector** corresponding to the eigenvalue $\lambda$ if $v \neq 0$ and $Av = \lambda v$.

$A \in M_{n\times n}(F)$ is **diagonalizable** over $F$ if there is $P \in M_{n\times n}(F)$ invertible such that $P^{-1}AP = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.

*Remark* 9.19. The word spectrum comes from quantum mechanics since the eigenvalues of the Hamiltonian operator are the energy of light emitted from a black body. So the study of eigenvalues and eigenvectors of a linear operator is also called spectral theory.

**Proposition 9.20.** *Let $A \in M_{n\times n}(F)$. Then $\lambda$ is an eigenvalue and $v$ is an eigenvector with eigenvalue $\lambda$ if and only if* $\det(\lambda I - A) = 0$ *and* $0 \neq v \in \ker(\lambda I - A)$.

*Proof.* Since $Av = \lambda v$, we have $(\lambda I - A)v = 0$ i.e. $v \in \ker(\lambda I - A)$. Since $v \neq 0$, $\lambda I - A$ is not invertible. Hence $\det(\lambda I - A) = 0$. ☐

**Definition 9.21.** Let $A \in M_{n\times n}(F)$. The polynomial $p_A(x) = \det(xI - A) \in F[x]$ is the **characteristic polynomial** of $A$. $\lambda$ is an eigenvalue of $A$ if and only if $\lambda$ is a root of $p_A(x)$.

The subspace $E(\lambda, A) = \ker(\lambda I - A)$ is called the **eigenspace** with eigenvalue $\lambda$.

The dimension $\dim E(\lambda, A) \ge 1$ is called the **geometric multiplicity** of $\lambda$.

The maximal integer $k$ such that $(x-\lambda)^k \mid p_A(x)$ is called the **algebraic multiplicity** of $\lambda$.

The following four examples are all possible situations that one may encounter when finding the eigenvalues and eigenvectors of a matrix.

**Example 9.22.** This is an example where the matrix is diagonalizable.

$A = \begin{pmatrix} 7 & -10 \\ 5 & -8 \end{pmatrix} \in M_{n\times n}(\mathbb{R})$

$p_A(x) = \det(xI - A) = \begin{pmatrix} x-7 & 10 \\ -5 & x+8 \end{pmatrix} = (x-7)(x+8) + 50 = x^2 + x - 6 = (x-2)(x+3)$.

The eigenvalues of $A$ are $\lambda_1 = 2$, $\lambda_2 = -3$ and both of algebraic multiplicity 1.

Solving $0 = (2I - A)v = \begin{pmatrix} -5 & 10 \\ -5 & 10 \end{pmatrix} v$ we have $v = c\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ for $c \in \mathbb{R}$.

Solving $0 = ((-3)I - A)v = \begin{pmatrix} -10 & 10 \\ -5 & 5 \end{pmatrix} v$ we have $v = c \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ for $c \in \mathbb{R}$.

Thus $\lambda_1 = 2$ is an eigenvalue with eigenvector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\lambda_2 = -3$ is another eigenvalue with eigenvector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. The geometric multiplicities of $\lambda_1, \lambda_2$ are both 1. If we write $P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$ then $AP = P \begin{pmatrix} 2 & \\ & -3 \end{pmatrix}$ i.e. $P^{-1}AP = \mathrm{diag}(2, -3)$.

**Example 9.23.** This is an example where $A$ is not diagonalizable since the field is too small to contain all eigenvalues of $A$.

$A = \begin{pmatrix} -1 & 2 \\ -1 & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{R})$

$p_A(x) = \det(xI - A) = \begin{pmatrix} x+1 & -2 \\ 1 & x-1 \end{pmatrix} = (x+1)(x-1) + 2 = x^2 + 1.$

There is no root of $p_A(x)$ in $\mathbb{R}$ which means $A$ has no eigenvalues and eigenvectors over $\mathbb{R}$. In particular, $A$ is not diagonalizable over $\mathbb{R}$.

**Example 9.24.** This is an example where we enlarge the field $F$ to contain all eigenvalues of $A$ and $A$ becomes diagonalizable in the enlarged field.

$A = \begin{pmatrix} -1 & 2 \\ -1 & 1 \end{pmatrix} \in M_{2\times 2}(\mathbb{C})$

$p_A(x) = \det(xI - A) = \begin{pmatrix} x+1 & -2 \\ 1 & x-1 \end{pmatrix} = (x+1)(x-1) + 2 = x^2 + 1 = (x-i)(x+i).$

The eigenvalues of $A$ in $\mathbb{C}$ are $\lambda_1 = i$, $\lambda_2 = -i$ both of algebraic multiplicity 1.

Solving $0 = (iI - A)v = \begin{pmatrix} i+1 & -2 \\ 1 & i-1 \end{pmatrix} v$ we have $v = c \begin{pmatrix} 1-i \\ 1 \end{pmatrix}$ for $c \in \mathbb{C}$.

Solving $0 = ((-i)I - A)v = \begin{pmatrix} -i+1 & -2 \\ 1 & -i-1 \end{pmatrix} v$ we have $v = c \begin{pmatrix} 1+i \\ 1 \end{pmatrix}$ for $c \in \mathbb{C}$.

Thus $\lambda_1 = i$ is an eigenvalue with eigenvector $\begin{pmatrix} 1-i \\ 1 \end{pmatrix}$ and $\lambda_2 = -i$ is another eigenvalue with eigenvector $\begin{pmatrix} 1+i \\ 1 \end{pmatrix}$. The geometric multiplicities of $\lambda_1 = i, \lambda_2 = -i$ are both 1. If we write $P = \begin{pmatrix} 1-i & 1+i \\ 1 & 1 \end{pmatrix}$ then $AP = P \begin{pmatrix} i & \\ & -i \end{pmatrix}$ i.e. $P^{-1}AP = \mathrm{diag}(i, -i)$.

*Remark* 9.25. Suppose we regard $A \in M_{n\times n}(\mathbb{R})$ as a matrix in $M_{n\times n}(\mathbb{C})$. If $\lambda \in \mathbb{C}$ is a complex eigenvalue with nonzero imaginary part, then $\bar{\lambda}$ is also an eigenvalue of $A$ by Lemma 9.16. In fact if $v$ is an eigenvector of $A$ with eigenvalue $\lambda$, then $\bar{v}$ is an eigenvector of $A$ with eigenvalue $\bar{\lambda}$. This is because if $Av = \lambda v$ then $A\bar{v} = \bar{A}\bar{v} = \overline{Av} = \overline{\lambda v} = \bar{\lambda}\bar{v}$. So in practice, we only need to compute half of the eigenvalues and eigenvectors.

*Remark* 9.26. In Example 9.23, $A$ is not diagonalizable over $\mathbb{R}$. In Example 9.24, we enlarge $\mathbb{R}$ to $\mathbb{C}$ and $A$ becomes diagonalizable. The price to pay is that the matrix $P$ is a complex matrix, not a real matrix. In general, the issue of the field $F$ not being large enough to contain all eigenvalues can be resolved by Proposition 9.17.

**Example 9.27.** This is an example where $A$ is not diagonalizable since there are not enough eigenvectors.

$A = \begin{pmatrix} -1 & -4 \\ 1 & 3 \end{pmatrix}$

$p_A(x) = \det(xI - A) = \begin{pmatrix} x+1 & 4 \\ -1 & x-3 \end{pmatrix} = (x+1)(x-3)+4 = x^2-2x+1 = (x-1)^2.$

$\lambda = 1$ is the only eigenvalue of $A$. It has algebraic multiplicity 2.

Solving $0 = (I - A)v = \begin{pmatrix} 2 & 4 \\ -1 & -2 \end{pmatrix} v$ we have $v = c\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ for $c \in \mathbb{R}$.

Thus $\lambda = 1$ is an eigenvalue with eigenvector $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$. Note that there is only 1 linearly independent eigenvector and the geometric multiplicity is 1.

*Remark* 9.28. In this example, $p_A(x)$ already splits in $\mathbb{R}[x]$, but there is not enough eigenvectors to diagonalize $A$ i.e. geometric multiplicity $<$ algebraic multiplicity. The indiagonalizability is essential and cannot be overcome by choosing a larger field. To see that $A$ is not diagonalizable from another perspective, suppose $A$ is diagonalizable. Then $P^{-1}AP = \mathrm{diag}(\lambda_1, \lambda_2)$. In this case, $p_A(x) = (x-1)^2$ and we must have $\lambda_1 = \lambda_2 = 1$. That means $P^{-1}AP = I$ and $A = PP^{-1} = I$ which is not possible.

## References


[1] Dummit, David Steven, and Richard M. Foote. Abstract algebra. Vol. 3. Hoboken: Wiley, 2004.